10-2015

Two formulas for success in social media: Learning and network effects

Liangfei QIU

Qian TANG
Singapore Management University, QIANTANG@smu.edu.sg

Andrew B. WHINSTON

DOI: https://doi.org/10.1080/07421222.2015.1138368

Two Formulas for Success in Social Media: Learning and Network Effects

Liangfei Qiu,[a] Qian Tang,[b] Andrew B. Whinston[c]

[a] Department of Information Systems and Operations Management
Warrington College of Business Administration, University of Florida
PO Box 117169, Gainesville, FL, USA 32611
Email: liangfei.qiu@warrington.ufl.edu

[b] School of Information Systems, Singapore Management University
80 Stamford Road, Singapore, Singapore 178902
Email: qiantang@smu.edu.sg

[c] McCombs School of Business, University of Texas at Austin
1 University Station, IROM, B6500, Austin, TX, USA 78712
Email: abw@uts.cc.utexas.edu

Published in Journal of Management Information Systems, 2015 October,
Volume 32, Issue 4, Pages 78-108
http://dx.doi.org/10.1080/07421222.2015.1138368

Two Formulas for Success in Social Media: Learning and Network Effects

Abstract

Recent years have witnessed an unprecedented explosion in information technology that enables dynamic diffusion of user-generated content in social networks. Online videos, in particular, have changed the landscape of marketing and entertainment, competing with premium content and spurring business innovations. In the present study, we examine how learning and network effects drive the diffusion of online videos. While learning happens through informational externalities, network effects are direct payoff externalities. Using a unique data set from YouTube, we empirically identify learning and network effects separately, and find that both mechanisms have statistically and economically significant effects on video views; furthermore, the mechanism that dominates depends on the video type. Specifically, although learning primarily drives the popularity of quality-oriented content, network effects make it also possible for attention-grabbing content to go viral. Theoretically, we show that, unlike the diffusion of movies, it is the combination of both learning and network effects that generate the multiplier effect for the diffusion of online videos. From a managerial perspective, providers can adopt different strategies to promote their videos accordingly, that is, signaling the quality or featuring the viewer base depending on the video type. Our results also suggest that YouTube can play a much greater role in encouraging the creation of original content by leveraging the multiplier effect.

Keywords: Learning, Network Effects, User-Generated Content, Social Contagion, Social Media

"You've got to create images they won't accept. Make them foam at the mouth. Force them to understand that they're living in a pretty queer world."

— Andre Malraux, *"Picasso Mask"* (1976), page 110.

Introduction

With new products such as consumer goods, financial services, and movies constantly flooding the markets, consumers face an already overwhelmingly large and rapidly growing choice set. Meanwhile, consumers receive various bits of information that generates two types of externalities: informational externalities and direct payoff externalities. Informational externalities exist when one's payoff depends on information that is privately held by others, and therefore are created when information about product quality is conveyed through direct communication/observation or indirect word-of-month [62, 51, 12, 16]. Positive payoff externalities exist when one's payoff depends positively on the number of other people who consume the product, and therefore are affected directly by others' actions [34]. Whereas the former is generally referred to as "learning" or "observational learning" [16, 26, 53], the latter is often called "network effects" or "network externalities" [34, 22, 18].

For products with strong network effects, creating a large user base is crucial in attracting new adopters. In contrast, generating positive word-of-mouth (WOM) is the key for products with prevailing learning effects. Susarla et al. [57] found that initial WOM conversations generated early in the life of a YouTube video have a persistent impact on subsequent popularity. In many situations, both mechanisms may be present at the same time. Which mechanism exists or dominates then depends on the specific product in question. When choosing a mobile network operator, network effects may dominate because of free mobile-to-mobile calling. When purchasing an HDTV, learning becomes the primary force because consumers are mainly concerned with quality.

Following the expectation–disconfirmation paradigm [44], we estimate the learning effects and network effects for social media content consumption by identifying disconfirmation/surprise resulting from a comparison of prior expectations with the actual consumption experience in the context of YouTube. According to this paradigm, expectation is the reference for a comparative judgment, and a negative (positive) surprise is rated below (above) this reference point [44]. Selecting online videos to watch is one of the most common choices consumers make every day. According to ComScore, the average user spent about 43 minutes watching online videos in June 2013, and Google websites (primarily YouTube) account for approximately 40% of that time, about 17 minutes.[1] According to YouTube statistics, 100 hours of video are uploaded to YouTube every minute. Given the vast reservoir of online videos, choosing videos to watch can become a complicated issue. On the one hand, consumers receive various pieces of information from friends, which can help them infer video quality. Such learning can take the form of direct communication and discussion with, or interpersonal and impersonal observation of, others [16]. On the other hand, frequent social sharing creates network effects when a video becomes a fad. Frequent social sharing may not be due to the quality of the video, but instead because of the emotional arousal while watching the video [35]. Particularly for a viral video, consumers have strong incentives to watch it so as to discuss it in social encounters. Many YouTube videos go viral with only pointless-seeming jokes, funny pictures, weird scenarios, or even offensive pranks, but they bring people together around something. Sharing allows people to engage with interesting people, including those they otherwise might not be aware of; furthermore, the way people interact with social media content is more about someone's emotional or even intellectual reaction [60]. The strength of learning depends on the surprise of the video quality and the accuracy of prior expectation, whereas the strength of network effects is related to social sharing which can be attributed to the emotional arousal generated. Therefore, the relative strengths of the two effects can vary across different types of videos.

---

[1] See http://www.comscore.com/insights/Press_Releases/2013/7/comScore_Releases_June_2013_US_Online_Video_Rankings.

Most existing studies on social contagion have focused on the Manski problem [38]: distinguishing general social contagion from homophily—the tendency of individuals to associate with similar others [8, 13, 61]. Few of them differentiated between the two mechanisms of social contagion: learning and network effects. To fill the gap, we are interested in the diffusion angle of learning and network effects, and study user-generated content in social media from the social contagion perspective. Figure 1 shows the conceptual framework of our study. Learning, the informational externalities, affects consumers through the quality information conveyed by peers, whereas network effects, the payoff externalities, influence consumers according to the size of the user base. Although these mechanisms lead to similar empirical outcome, their implications are vastly different. If contagion is generated mainly by network effects, then seeding strategies, which determine the initial set of targeted consumers, will by implication have a strong influence on the success of viral marketing. Accordingly, a firm can amplify social contagion and accelerate product purchases by offering introductory discounts [31]. If learning is the dominant effect, however, seeding will not be effective unless the initial consumers generate positive word-of-mouth. Consumers can infer that the high demand among their peers is caused by the introductory discount rather than the high product quality [46]. Both cases are theoretically plausible and need to be empirically distinguished.

Given the lack of pre-release marketing effort, these two mechanisms of social contagion are particularly important for user-generated content (UGC). To the best of our knowledge, our paper is the first to disentangle consumer learning and network effects in the context of online video sharing. Previous studies on online video sharing using traditional Bass model to study the diffusion process have shown that the existing user base has a positive effect on future adoption but make no distinction between learning and network effects regarding the underlying mechanisms [56, 58]. To fill the gap, we differentiate the two mechanisms by examining how video consumers react to different types of information. Additionally, we categorize popular videos into high-quality videos and attention-grabbing videos according to the consistency in their ratings, and examine the different impacts of the two mechanisms in their diffusion process respectively. Our empirical results suggest that both mechanisms affect the diffusion of social media content significantly, with consumer learning having a greater influence on high-quality content, and network effects having a greater influence on attention-grabbing content. The implications derived from studying YouTube can carry over to other consumer choice problems as well.

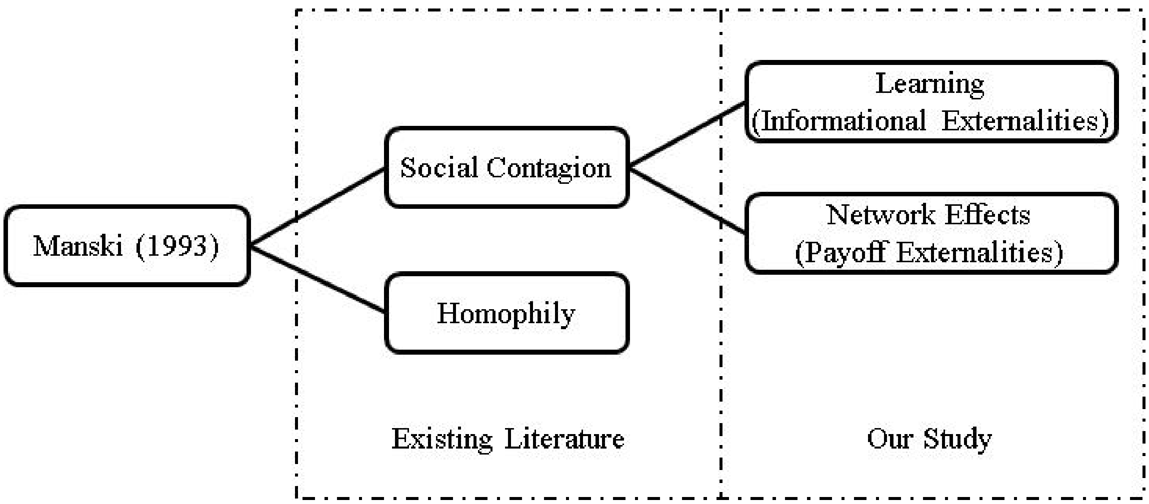

Figure 1. The Conceptual Framework of Learning and Network Effects

Literature Review

Identification of social contagion has long been recognized as an econometric challenge [38]. Failure to account for contextual effects or homophily often leads to an overestimation of the effect of social contagion. Aral et al. [7] distinguished influence-based contagion from homophily-driven diffusion using a dynamic matched sample of global instant messaging users. Within the framework of social contagion, studies have focused on distinguishing learning from other contagion mechanisms such as saliency effect (i.e., observed choices are more salient than alternative choices) and conformity concerns (i.e., the social pressure to adopt the choice made by the majority). Cai et al. [19] used a field experiment to show that observational learning, rather than saliency effect, affects customers' choices. Shi and Whinston [52] studied observational learning in the context of location-based networks.

Network effects, or network externalities, have been widely studied as an important driver for the diffusion of technology products or services, such as standards [22], software [27], and social networks [33, 64, 47]. These products or services become more valuable as more people use them. In IT and electronic commerce areas, network externalities may develop from exchange, stability, or extrinsic benefits [27]. For the diffusion of user-generated content in social media, network effects mainly result from exchange; that is, each new content consumer adds potential value through exchange with other consumers [24].

While learning is generated by information externalities, network effects are created because of payoff externalities. In spite of this fundamental difference, they often coexist and even interact with each other in the diffusion of products or innovations [21, 23]. Rather than distinguishing informational externalities from payoff externalities, most studies have focused on using the coexistence of both to explain herding or informational cascades. Bikhchandami et al. [16] argued that an informational cascade resulting from observational learning is very fragile to small shocks, whereas the uniformity created by network externalities does not display the fragility. Moretti [41] showed that social learning is a more important determinant of sales in the movie industry than network effects. However, we suspect that the same can apply to online videos. Consumers have more precise prior information such as crew information, movie trailers, critics, and so on, to estimate the quality of a movie, but much less information for social media videos, considering the limited marketing campaigns and alternative information sources of user-generated content. Therefore, it is unlikely that a movie can go viral solely because of network effects, but social media videos might.

In this study, we adopt the expectation–disconfirmation paradigm in marketing literature [44]. Based on this paradigm, customer satisfaction has three main antecedents: prior expectations, ex-post quality, and disconfirmation/surprise. We introduce these three key constructs in the analysis of online videos. In consumer satisfaction literature, prior expectations have been conceptualized as the beliefs about a product formed by consumers' prior experiences and exposure to firms' marketing efforts [43, 44]. In our context of YouTube videos, when consumers consider watching a video, they utilize ex-ante information, such as video providers' reputation, to form their expectations of the underlying video quality. Ex-post quality is defined as consumers' post-usage evaluation about how a product fulfills their needs, wants, and desires [39]. In Susarla et al. [55], the perception of an application service provider's service quality results from an ex-post evaluation of the service. Disconfirmation/surprise is defined as discrepancies between consumers' expectations and the ex-post quality [44]. Consumers' expectations will be negatively disconfirmed if the product performs worse than expected and positively disconfirmed if performance is better than anticipated. We use this paradigm to study consumer learning, by which consumers use information to infer product quality and act accordingly [59]. The positive (negative) disconfirmation thus leads to greater positive (negative) word of mouth and more (fewer) people watching the video subsequently.

Hypothesis Development

A Simple Analytical Model

YouTube videos are experience goods whose quality cannot be fully observed by consumers ex-ante but can be ascertained upon consumption. Therefore, before consumption, consumers are never certain about the quality, but can always acquire useful information from friends who have already watched the videos. If the revealed information suggests that the true video quality is higher than the expected value, we call it positive disconfirmation/surprise. If the information suggests that the true quality is lower than the expected value, we call it negative disconfirmation/surprise.

Learning is a process of adjusting beliefs about the quality according to disconfirmation. We build a simple analytical model of individual belief updating that can capture the underlying learning process discussed in our study. The literature on observational learning [12] examines the learning that occurs through observing other people's behaviors. We consider a general learning model through direct communication and discussion with, and observation of others. We capture the learning process with a Bayesian learning model, where each consumer receives feedback from peers and updates the prior belief of the video quality.[2] The utility that a representative consumer i obtains from watching a YouTube video j is given by: $u_{ij} = V_j + \eta_{ij}, \eta_{ij} \sim N(0, 1/\rho_\eta)$, where V_j is the latent quality of the video, and η_{ij} represents the unobserved taste heterogeneity. Each individual knows her personal taste, η_{ij}. At time 1, video j is posted on YouTube. Since there is no specific prior information pertaining to the video before it is posted, consumers share a common prior on the quality of the video, given by $V_j \sim N\left(X_j'\beta, 1/\rho_{V_j}\right)$, where X_j is a vector of the observable characteristics of video j's provider before watching and the initial characteristics of video j described in Table 3, $X_j'\beta$ is the ex-ante expectation of quality, and ρ_{V_j} is the precision of prior for video j. Notice that the video is newly published, so no learning occurs at time 1. A consumer chooses to watch it if the ex-ante expected utility is no less than the opportunity cost of watching video j, c_{ij}:

$$E_1[u_{ij}|I_1] = X_j'\beta + \eta_{ij} \geq c_{ij},$$

where $E_t[u_{ij}|I_t]$ represents consumer i's expected utility of video j at time t given the information set at time t, I_t. Accordingly, the probability that a consumer watches video j at time 1 is:

$$\Pi_{i1} = Pr\big(E_1[u_{ij}|I_1] \geq c_{ij}\big) = \Phi\big[\sqrt{\rho_\eta}(X_j'\beta - c_{ij})\big].$$

With learning, consumers have more information at time 2 because they receive feedback from friends. We assume that a consumer i at time 2 has k friends. Among them, k_1 friends have watched the video at time 1. These friends communicate to consumer i their ex-post utilities after watching the video, u_{mj}, where m =1, 2, 3, ... , k_1.

As a result, at time 2, consumer i's information set consists of the ex-post utilities of some friends and the number of friends who have not watched the video. Consumer i at time 2 estimates the

[2] Following Banerjee [12], the timing of consumption is exogenously given, and we do not consider the strategically behavior of delaying the decision making process to obtain more feedback.

quality by maximizing the likelihood of the observed evidence:

$$L\left[u_{mj}, m = 1, 2, 3, \dots, k_1; k - k_1 \middle| V_j\right]$$

$$= \Pi_{m=1}^{k_1} f\left(u_{mj}\right) \cdot \Pi_{m=k_1+1}^{k} \Pr\left(E_1\left[u_{mj} \middle| I_1\right] < c_{ij}\right),$$

where $f\left(u_{mj}\right)$ is the likelihood of observing u_{mj}. The maximum likelihood estimator, G_{ij2}, is an estimate of V_j. It is unbiased and asymptotically normal:

$$G_{ij2} \sim N\left(V_j, 1/d_{ij2}\right),$$

where $d_{ij2} = -E\left[\left(\frac{\partial lnL}{\partial V_j}\right)^2\right]$ [5].

Consumers update the prior according to Bayes' rule. At time 2, her expected utility becomes the weighted average of the prior mean and the maximum likelihood estimator:

$$E_2\left[u_{ij} \middle| I_2\right] = \frac{\rho_{V_j}}{\rho_{V_j} + d_{ij2}} X_j' \beta + \frac{d_{ij2}}{\rho_{V_j} + d_{ij2}} G_{ij2} + \eta_{ij}.$$

Note that as time goes on, consumers place less weight on the prior mean. Because consumers receive more information at time 2, the prior becomes a less important factor in the decision making process. The probability that consumer i watches video j at time 2 is:

$$\Pi_{i2} = Pr\left(E_2[u_{ij}|I_2] \geq c_{ij}\right) = \Phi\left(\frac{\frac{\rho_{V_j}}{\rho_{V_j}+d_{ij2}} X_j'\beta + \frac{d_{ij2}}{\rho_{V_j}+d_{ij2}} V_j - c_{ij}}{\sqrt{d_{ij2}/\left(\rho_{V_j} + d_{ij2}\right)^2 + 1/\rho_\eta}}\right).$$

The decision making process proceeds in the same way at time *T*. Consumer i has k_t friends who decide to watch the video at time t, where $t = 1, 2, 3, \dots, T-1$. The probability that consumer i watches video j at time T is:

$$\Pi_{iT} = Pr\left(E_T[u_{ij}|I_T] \geq c_{ij}\right) = \Phi\left[\frac{\alpha_T X_j'\beta + (1-\alpha_T)V_j - c_{ij}}{g_T}\right].$$

where $g_T = \sqrt{\left(\sum_{t=2}^{T} d_{it}\right)/\left(\rho_{V_j} + \sum_{t=2}^{T} d_{ijt}\right)^2 + 1/\rho_\eta}$, and $\alpha_T = \frac{\rho_{V_j}}{\rho_{V_j} + \sum_{t=2}^{T} d_{ijt}}$.

It is evident that α_T decreases with T. As time T grows, the probability of watching videos relies less on the ex-ante prior and more on learning. In the analytical model, If the revealed quality of the video is higher than the mean of the ex-ante prior, $V_j - X_j'\beta > 0$, we call it positive disconfirmation/surprise. If the revealed quality of the video is lower than the prior mean, $V_j - X_j'\beta < 0$, we call it negative disconfirmation/surprise. We empirically operationalize disconfirmation/surprise as the difference between realized video views and predicted video views at time 1. This empirical construct corresponds to $V_j - X_j'\beta$ in our analytical model.

In essence, learning is a process of adjusting beliefs about the quality according to disconfirmation. We obtain the following two results, and the proofs can be found in Appendix A.

Result 1: If positive/negative disconfirmation is sufficiently large, then Π_{iT} is increasing/decreasing in T.

Result 2: The distance between the true video quality and consumers' expected video quality, $|V_j - E_T[V_j|I_T]|$, is decreasing in T.

Hypotheses of Learning

Intuitively, if consumers learn about disconfirmation over time, a video that has positive disconfirmation will have a higher growth rate of viewership than a video that has negative disconfirmation over time. The expectancy disconfirmation plays a major role in determining satisfaction [44]. Satisfaction or dissatisfaction with consumption experience is generally regarded as the key antecedent of product-related word of mouth [6]. As dissatisfaction increases, the tendency of consumers to engage in negative word-of-mouth activities increases [48]. In our context, the positive disconfirmation/surprise increases consumer satisfaction, and then leads consumers to engage in greater positive word of mouth. Therefore, the growth rate of viewership increases more over time in the case of a positive surprise than that of a negative one, which is consistent with Results 1 and 2.

Second, if learning is significant, we should observe that the magnitude of learning is moderated by the quality uncertainty of videos. The intuition is that the learning effect should be more pronounced for videos with less-precise consumer priors. If a consumer is very uncertain about the quality of a video, the impact of consumer learning is large: The additional information she learns from her friends should be more valuable than when she knows the quality precisely. Thus, learning should be stronger for videos that are less familiar to consumers. Combining the two predictions from learning, we have the following hypothesis:

Hypothesis 1. *In the presence of learning (informational externalities), (i) a video that has positive disconfirmation would have a higher growth rate of viewership than a video that has negative disconfirmation over time. (ii) Furthermore, positive disconfirmation has a greater impact on videos with less-precise priors.*

However, network effects, or network externalities, occur if the consumers derive utility directly from the consumption by others despite of the quality of the product. We test for the existence of network effects by examining whether consumers respond to disconfirmation that is irrelevant to video quality. If only learning exists, such negative or positive disconfirmation will have no significant impact on future views because it does not reveal the video quality. However, if network effects also exist, a lower-than-expected (higher-than-expected) views caused by non-quality factors, can still reduce (increase) future video views and trigger the feedback loop. In other words, the lower-than-expected (high-than-expected) views would lead to potential viewers being less (more) likely to watch the video. Therefore, the growth rate of viewership would be affected accordingly over time. This leads to the following hypothesis:

Hypothesis 2. *In the presence of network effects (payoff externalities), a video that has positive disconfirmation that is unrelated to video quality, would have a higher growth rate of viewership than a video that has negative disconfirmation that is unrelated to video quality, over time.*

Both learning and network effects may exist on YouTube. Therefore, a video can go viral in either way. Lee and Raghu [36] showed that the sales performance of an app depends on several important attributes that are related to learning and network effects. Through examining the most popular videos on YouTube, we can categorize them into two distinct groups: one group consists of videos that feature high quality, engaging scenes, and articulated story lines (high-quality videos), and the other group of videos often includes questionable behaviors that deviate from social norms yet still gains tremendous popularity (attention grabbers). The "Pussy Riot" incident in Russia serves as a good example of a typical attention grabber. This Russia-based feminist rock band protested against the political scene in Russia through unorthodox musical performances and produced YouTube videos that went viral overnight. It is worth noting that Pussy Riot did not gain international fame through their musicality per se; instead, most viewers were drawn to those videos out of curiosity and were interested in the messages the music carried.

A strategy often adopted by attention grabbers is the inclusion of controversial elements in videos. Such instances often provoke controversy and stir heated discussion revolving around those contents. In contrast to those quality-oriented productions, the goal of attention grabbers is primarily to attract attention or promote ideas. Intuitively speaking, we would not expect too much learning effect to take place for the popularity of this type of video. In an analytical model, Eliaz and Spiegler [25] showed that a firm can earn higher profits by employing pure attention grabbers with positive probability. Similarly, we propose that, as suggested by their discussion-provoking nature, videos with attention-grabbing content can initiate higher network effects, and viewers find it valuable to watch them because these videos allow them to engage in discussions with their social contacts and generate more social sharing. Therefore, we hypothesize that attention-grabbing videos gains popularity mostly through network effects as opposed to learning, and the opposite is true for high-quality videos:

Hypothesis 3. *(a) Network effects (payoff externalities) are more pronounced for videos with attention-grabbing content. (b) Learning (informational externalities) is more pronounced for high-quality videos.*

Data

We collected data on newly posted videos from YouTube. As the world's largest video viewing and sharing website, YouTube has enormous numbers of videos, which makes random sampling infeasible. Instead, we focus on the most active providers by selecting the top 1,000 YouTube providers (in terms of total video views) identified for June 2011.[3] We collected a daily panel of data on these providers and their new videos for one month, from March 1, 2012 to March 31, 2012. Our sample includes 302 new videos published by top providers on March 1, 2012. We use one day as the time unit of analysis to capture the fast-changing nature of online videos.

The provider level data include provider ID, data collection date, date when the provider joined YouTube, number of subscribers to the provider's channel, total views of all the provider's videos, total views of the provider's channel page, number of videos, number of friends, number of subscriptions the provider has to other providers, channel views rank, and video views rank. The video level data include video ID, data collection date, date when the video is posted, the provider of the video, number of views, category in which the video belongs, video length, whether the video has an in-stream ad, average rating, number of times the video is favorited by viewers (number of favorites), and number of comments. All videos in our sample were published on March 1, 2012. Because YouTube Analytics data is updated daily, the first day in our analysis is March 2, 2012. Summary statistics of the video characteristics at the beginning of our data collection period are reported in Table 1. We assume that each viewer watches a video only once. Although consumers may repeatedly watch a video, the bias caused by repeated viewings is small if logs of views are used instead of views [56]. Table 2 provides summary statistics of the characteristics of our YouTube providers.

[3] To make our results generalizable to unpopular providers and to compare with top providers, we also identified 2,236 new providers who joined YouTube during January 2012. The basic results are robust.

Table 1.The First-Day Video Characteristics

Variables	Mean	Std. Dev.	Min	Max
Number of video views	1,717	6,882	1	107,628
Video rating	4.73	0.51	1	5
Number of favorites	89	399	0	6,800
Number of comments	245	838	0	9,832
In-stream ads (yes-1, no-0)	0.34	0.48	0	1

Some YouTube providers also post their video links on Twitter. We control for these personal marketing efforts when estimating social contagion. Using Twitter application programming interface (API), we collected a random selected sample of all Twitter data containing the key word #YouTube or video. Using the collected tweets, we analyzed the included shortened uniform resource locator (URL) link related to YouTube and recovered the unique YouTube video ID. Then, we used a simple natural language processing on tweet content to identify the tweets posted by video providers as new video announcements, such as "I uploaded a new video … please check out." For our sample videos, we find that around 5% of top providers did self-promotion for their new videos on Twitter.

Table 2. The First-Day Chanel Characteristics

Variables	Mean	Std. Dev.	Min	Max
Total channel views	8,292,525	1.72e+07	214	1.75e+08
Total video views	1.17e+08	2.09e+08	997	1.55e+09
Number of subscribers	187,852	392,605	51	5,109,145
Number of subscriptions	131	848	0	17,641
Number of videos	184	257	1	969
Number of friends	10,707	19,954	0	120,570
Number of channel favorites	91	263	0	3,292

Empirical Framework

Identification of the Disconfirmation/Surprise

Following Barro [14]; Hirshleifer et al. [30]; and Moretti [41], we empirically operationalize disconfirmation/surprise as the difference between realized video views and predicted video views at time 1 (March 2, 2012). More specifically, we argue that realized video views can be predicted by the information set of viewers before time 1, reflecting the ex-ante rational expectation of video quality.

The viewer's information set includes two pieces of information: (1) the characteristics of YouTube providers (channels) before time 1, including the log of total views of channel *j*'s videos, lvviews; the log of total views of the provider's channel page, lcviews; the log number of uploaded videos of the channel, lvideos; the log number of the provider's subscribers, lsubs; the log number of other providers the provider subscribes to, lsubscriptions; the log number of times a channel is favorited by viewers, lchannel_favs, the average rating of the channel for all its videos, avg_chanel_rating; the variance of the ratings of this channel's videos, var_chanel_rating; and the age of the channel (in terms of days), channel_age. The characteristics of YouTube providers are reasonable measures of expected video quality. YouTube allows consumers to subscribe to the providers they would like to follow. By subscribing to a provider, they are informed immediately whenever the provider posts a new video. The prior information about the provider shapes to a large extent a consumer's expectation of the provider's new videos. (2) The initial characteristics of a new video. When a new video is posted, a viewer can infer the video quality from some initial video characteristics, including a set of dummy variables indicating the video category, a dummy variable indicating whether the video was self-promoted by the provider on Twitter, tweet_upload, and a dummy variable, length, which takes the value 1 if the video length is longer than 10 minutes, and 0 otherwise.

Table 3. Identification of Disconfirmation: First-Stage Regression

	(1) OLS	(2) OLS	(3) OLS	(4) Rating as Surprise
lvviews	0.497***	0.493***	0.498***	0.0742***
	[9.714]	[10.27]	[10.49]	[3.024]
lcviews	0.259***	0.253***	0.249***	0.0451*
	[5.948]	[6.374]	[6.367]	[1.729]
lvideos	0.0180	0.00926	0.0121	0.0331***
	[0.883]	[0.492]	[0.655]	[2.832]
lsubs	-0.0707*	-0.0662*	-0.0667*	-0.0176
	[-1.890]	[-1.825]	[-1.844]	[-0.825]
lsubscriptions	-0.0194	-0.0186	-0.0194	-0.0275
	[-0.652]	[-0.672]	[-0.705]	[-1.625]
lchannel_favs	0.0200	0.0366	0.0380	0.0297*
	[0.653]	[1.319]	[1.373]	[1.713]

length	-0.0870	-0.0849		0.123
	[-0.657]	[-0.729]		[1.636]
avg_chanel_rating	0.235**	0.268**	0.266**	0.102**
	[2.186]	[2.552]	[2.542]	[2.278]
var_chanel_rating	-0.352*	-0.373*	-0.375*	-0.515***
	[-1.730]	[-1.908]	[-1.922]	[-4.480]
channel_age	-0.000119	-0.000123	-9.92e-05	-0.000127
	[-1.107]	[-1.291]	[-1.095]	[1.077]
tweet_upload	-0.185	-0.143		0.129
	[-0.740]	[-0.591]		[0.917]
video category dummies	Y	N	N	Y
Constant	-6.523***	-6.984***	-7.057***	5.090***
	[-5.637]	[-11.11]	[-11.39]	[5.886]
Observations	302	302	302	302
R-squared	0.674	0.661	0.661	0.704

t-statistics in brackets: *** p<0.01, ** p<0.05, * p<0.1

We use the residual from a regression of the log of the first-day views on the characteristics of YouTube providers and the initial characteristics of a new video as a measure of disconfirmation/surprise. More specifically, the predicted value of the dependent variable from the estimated linear equation including all independent variables is the rational expectation of the dependent variable (the log of the first-day views). The residual is the deviation of the actual video views from the rational expectation. The literature on rational expectations [42, 50, 14] asserts that on average, the deviation from rational expectation should be zero. Table 3 presents the first-stage regression results. The residual is used as a measure for disconfirmation/surprise. The correlation among independent variables and the variance inflation factors (VIF) are shown in Tables B.1 and B.2 in Appendix B. Columns 1, 2, and 3 in Table 3 show the regression results under different regression specifications (column 1: full regression model, column 2: without video category dummies; column 3: including only the characteristics of YouTube providers), with column 4 using another empirical measure of surprises: the difference between realized video ratings and predicted video ratings at time 1. The results are consistent across these four different model specifications. As an additional check, we also construct a measure of disconfirmation using the difference between realized views and predicted views in the first week instead of that on the first day, and the result is robust.

A Test of Consumer Learning

Figure 2 shows a clear illustrative example of videos with different disconfirmation/surprises. The figure plots the daily video views for a video with a positive disconfirmation/surprise (video 2) and a video with a negative disconfirmation/surprise (video 1). These two videos belong to the same YouTube video category and have similar initial views, but experience different growth patterns: Video 2, having positive disconfirmation/surprise, has a significantly higher growth rate than video 1, having a negative disconfirmation/surprise. The first-day views of video 1 and video 2 are roughly the same (304 and 449 respectively). However, at the end of our sample period, views of video 1 and views of video 2 are $1{,}102$ and $25{,}508$ respectively. This striking difference is likely to be caused by social learning over time.

To formally test whether the difference between videos with positive disconfirmation and those with negative disconfirmation is statistically significant, we estimate the following model:

$$\ln views_{j,t} = b_0 + b_1 t + a_j + \Psi'_{j,t} b_2 + b_3\left(t \times D_j\right) + b_4\left(t \times EV_j\right) + \mu_{j,t}, \tag{1}$$

where $\ln views_{j,t}$ is the log of views of video j at time t, t represents the time period, a_j represents unobserved fixed effect of video j, and $\Psi_{j,t}$ includes the characteristics of video j that change over time, such as $rating$ (the video ratings), $favs$ (the number of YouTube Favorites), $comment$ (the number of video comments), $lvideo$ (the log number of uploaded videos of the channel), $lvviews$ (the log of total video views of the channel), and $subs$ (the number of channel subscriptions). We control for the marketing efforts of YouTube providers on Twitter measured by sum_upload_{jt}, the total number of tweets containing the unique YouTube video ID and the word "uploaded". EV_j is the expected video quality measured by the predicted value of the dependent variable obtained from the first-stage regression described in Table 3. In the regression, D_j is a dummy variable indicating whether the disconfirmation/surprise of video j is positive, and $\mu_{j,t}$ is the error term. We expect that b_4 is not significantly different from 0 because the expected quality should not change the growth rate of video views over time after controlling for other variables in Model (1).

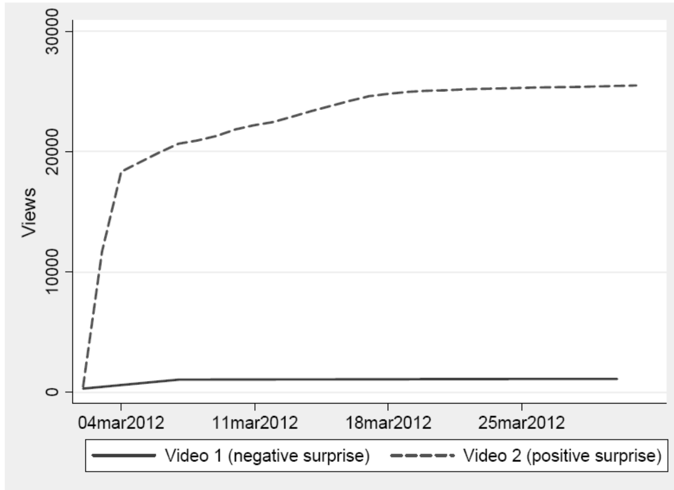

Figure 2. Daily Views for Videos with Different Disconfirmation

Following the literature on treatment effects (Wooldridge 2002), we make the unconfoundedness assumption: $t \times D_j$ is strictly exogenous. Note that the correlation between $t \times D_j$ and $\mu_{j,r}$ for any time t and time r causes inconsistency in regression coefficients. Thus, we need to control for the time-varying heterogeneity ($\Psi_{j,t}$), and the unobserved fixed effects in the regression. If the disconfirmation assignment (positive or negative) changes in reaction to past outcomes on $\ln views_{j,t}$, strict exogeneity can be violated [63]. However, because the surprise assignment is determined at time 1 and is independent of the idiosyncratic views shocks in period t, strict exogeneity is a reasonable assumption.

We are interested in b_3. In the presence of learning, viewership trends for positive and negative disconfirmation videos should diverge over time. The coefficient b_3 captures the size of this divergence caused by learning. If $b_3 > 0$, then the difference between the growth rates is positive, supporting Hypothesis 1(i). If $b_3 = 0$, then the growth rates of video views with different surprises are the same, which indicates no significant learning.

The fixed effects regression results are shown in Table 4. Column 1 shows the results from a regression that includes all the coefficients specified in equation (1) except b_4. The cluster robust t-statistics are shown in brackets. In this regression, $\hat{b}_3$ is significantly positive, which confirms Hypothesis 1 (i). In the regression model, after controlling for video rating, favorites, and comments, the impact of a positive disconfirmation is still significantly positive. The coefficient $\hat{b}_3$ is also practically significant. The effect of positive disconfirmation is considerable: A video that has positive disconfirmation can experience about 5% higher growth rate of viewership every day than a video that has a negative surprise. It is also worth noting that the coefficient b_3 reported in Table 4 is the *upper bound* of the learning effect, and we cannot rule out other explanations, such as network effects. As expected, Column 2 shows that $\hat{b}_4$ is insignificant, which implies that the anticipated quality does not have a significant impact on the growth rate of video views. This result is reminiscent of rational expectations models: Only unanticipated factors affect real economic variables [14].

Table 4. Fixed Effects Regression of Video Views on Disconfirmation/Surprises: A Test of Learning

	(1) FE	(2) FE	(3) Bootstrap	(4) Tech/Music	(5) Rating as Surprise
$(\boldsymbol{t} \times \boldsymbol{D_j})$	0.0503***	0.0521***	0.0503***	0.0516***	0.0324***
	[25.60]	[21.76]	[21.23]	[24.91]	[22.45]
$\boldsymbol{t}$	0.0424***	0.0331***	0.0424***	0.0402***	0.0213***
	[13.77]	[10.24]	[19.31]	[13.81]	[11.45]
rating	0.0140	0.0237	0.0140	0.0137	0.0425
	[0.0210]	[0.0728]	[0.0222]	[0.0205]	[0.148]
comment	0.000838***	0.000822***	0.000838***	0.000832***	0.000572***
	[3.655]	[4.040]	[5.334]	[3.658]	[3.253]
favs	3.13e-05	-0.000116	3.13e-05	3.07e-05	0.000148
	[0.116]	[-0.542]	[0.141]	[0.114]	[0.415]
sum_upload	0.507***	0.390***	0.507	0.502***	0.321***
	[2.844]	[2.984]	[0.926]	[2.845]	[3.027]
$(\boldsymbol{t} \times \boldsymbol{EV_j})$		0.00670			
		[0.374]			
$(\boldsymbol{t} \times \boldsymbol{D_j} \times \boldsymbol{Tech})$				0.0236***	
				[3.192]	
$(\boldsymbol{t} \times \boldsymbol{D_j} \times \boldsymbol{Music})$				-0.0221***	
				[-3.349]	
Observations	9060	9060	9060	9060	9060
R-squared	0.286	0.283	0.286	0.289	0.235

Robust t-statistics in brackets: *** p<0.01, ** p<0.05, * p<0.1

Using residuals to measure unanticipated variables (disconfirmation) has a long tradition in macroeconomics and finance. A number of studies use the following two-step regression procedure to estimate the effect of the unanticipated variables [50, 14, 30, 41]: First, the residuals from a separate auxiliary regression are used as a proxy for the unanticipated variable, and then the residuals are used as an explanatory variable in the equation of interest [11]. Simply using ordinary least squares (OLS) without adjusting for the extra variance of the generated regressor term (the surprise) can yield consistent estimates but invalid statistical inferences [45, 63]. It is crucial to address this issue in the context of social media because the data tend to be noisier. We use the two-step bootstrapping method proposed by Cameron and Trivedi [20] to obtain proper statistical inferences. The asymptotically refined result is presented in Column 3.

As an additional robustness check, we add the lagged dependent variable as an independent variable and estimate the following dynamic panel data model:

$$\ln views_{j,t} = b_0 + b_1 t + a_j + \Psi'_{j,t} b_2 + b_3 \left(t \times D_j\right) + b_4 \ln views_{j,t-1} + \mu_{j,t}.$$

In a dynamic panel data model, a well-known "dynamic panel bias" could be caused by the fact that the lagged dependent variables are correlated with the transformed error term. In order to remove this bias, we report the Arellano–Bond estimator [9] by using the difference generalized method of moments (difference GMM) in column 1 of Table 5. After controlling for the log of video views lagged 1 period, $\ln views_{j,t-1}$, the coefficient on the interaction term, $(t \times D_j)$, is still significantly positive, which supports Hypothesis 1 (a). A potential weakness in the Arellano–Bond estimator is that difference GMM performs poorly when past levels of the dependent variable convey little information about future changes. The Arellano–Bover/Blundell–Bond estimator [10, 17] allows the introduction of more instruments and can dramatically improve efficiency. The results are consistent and shown in column 2 of Table 5. After the regression, we also perform Arellano–Bond test for autocorrelation in a panel: No evidence shows that the lagged instruments are invalid.

Table 5. Dynamic Panel-Data Estimators of Video Views on Disconfirmation/Surprises

	(1) Arellano–Bond estimation	(2) Arellano-Bover/Blundell-Bond estimation
$\ln views_{j,t-1}$	0.213***	0.456***
	[6.072]	[7.133]
$(t \times D_j)$	0.0402**	0.0413***
	[2.426]	[3.602]
t	0.0233***	0.0201***
	[5.692]	[6.074]
rating	0.648*	0.576***
	[1.930]	[3.425]
comment	0.000259***	0.000186
	[2.746]	[1.060]
favs	0.000135	-3.81e-05
	[1.273]	[-0.118]
sum_upload	0.329***	0.0430
	[2.669]	[0.126]
Observations	9060	9060

Robust t-statistics in brackets: *** p<0.01, ** p<0.05, * p<0.1

Learning effects may differ across video categories. We focus on two specific video categories on YouTube: "tech" and "music." According to a survey by Sysomos Inc. (http://www.sysomos.com/reports/youtube/), music is the most popular category on YouTube, and tech is the least popular category. We estimate the following regression model in column 4 of Table 4:

$$\ln views_{j,t} = b_0 + b_1 t + a_j + \Psi'_{j,t} b_2 + b_3\left(t \times D_j\right) + b_4\left(t \times D_j \times Tech_j\right)$$

$$+b_5\left(t \times D_j \times Music_j\right) + \mu_{j,t},$$

where $Tech_j$ is a dummy variable that takes the value 1 if video j belongs to "tech" category, and 0 otherwise, and $Music_j$ is a dummy variable that takes the value 1 if video j belongs to "music" category, and 0 otherwise. We find that the coefficient on $t \times D_j \times Tech_j$ is significantly positive, but the coefficient on $t \times D_j \times Music_j$ is significantly negative. This finding indicates that learning affects videos of unpopular categories more. The implication is that the role of learning becomes more salient in a niche market than in a mass market.[4] Column 5 shows that the results are robust when we define the surprise as the difference between realized video ratings and predicted video ratings.

[4] An alternative explanation is that, in contrast with a tech video, consumers are more likely to watch a music video without asking for others' opinions because once they find it is a good music video, they can watch it more than once in future.

Hypothesis 1 (ii) indicates that learning is more important for videos with less-precise priors. To test this prediction, we estimate the following model:

$$\ln views_{j,t} = b_0 + b_1 t + a_j + \Psi'_{jt} b_2 + b_3(t \times D_j)$$
$$+b_4(t \times prior_j) + b_5(t \times D_j \times prior_j) + \mu_{jt},$$

where $prior_j$ is a measure of the prior precision of video j. Here we propose using the variance of the ratings of video j's provider/channel's existing videos at time 1, var_chanel_rating, to measure prior precision. Viewers are more certain about the quality of the provider/channel and thus his/her new videos if the ratings for his/her existing videos are more consistent. So it is reasonable to assume that the higher the variance in ratings of existing videos, the less precise the viewers' priors. To operationalize, we define $prior = 1/\text{var\_chanel\_rating}$.

Table 6 shows that the empirical results support Hypothesis 1 (ii). The coefficient on $(\boldsymbol{t} \times \boldsymbol{D_j} \times \boldsymbol{prior_j})$ is significantly negative. To interpret, we can consider two identical videos with the same positive disconfirmation except for the fact that first one was posted by a provider with a higher value of $1/\text{var\_chanel\_rating}$. In the presence of learning, a negative b_5 leads to a lower growth rate of views for the first video, as a result of learning from surprises. In other words, learning has a greater effect on videos with less-precise priors. The size of the coefficient on $(\boldsymbol{t} \times \boldsymbol{D_j} \times \boldsymbol{prior_j})$ is also considerable: If the value of $1/\text{var\_chanel\_rating}$ of the first video is one standard deviation higher than that of the second video, the daily growth rate difference is about 1.4%. Note that this moderating role of quality uncertainty on learning is a unique feature of learning, and cannot be explained by network effects. Our empirical results reported in Table 6 further confirm the existence of learning.

Table 6. The Effect of Prior Precision on Learning

	(1) FE
$(t \times D_j)$	0.0490***
	[18.76]
$(t \times prior_j)$	3.10e-05***
	[2.642]
$(t \times D_j \times prior_j)$	-3.43e-05***
	[-2.757]
t	0.0410***
	[11.96]
Rating	0.142
	[0.206]
Comment	0.00102***
	[3.621]
Favs	0.000206
	[0.615]
sum_upload	0.497***
	[3.005]
Observations	9060
R-squared	0.324

Robust t-statistics in brackets: *** p<0.01, ** p<0.05, * p<0.1

A striking pattern in the data is that video views are remarkably skewed. For example, the top 10 videos account for 47.46% of total views, and the top 30 account for 66.81% in our sample. Quantile regression analysis is particularly useful when the conditional distribution of video views is heterogeneous and does not have a "standard" shape. The quantile regression models allow us to account for unobserved heterogeneity and heterogeneous covariates effects. We can examine the effect of disconfirmation on the entire viewership distribution instead of the conditional mean of viewership distribution by using quantile regression and gain a better understanding on viewership inequality. A simple differencing strategy used in fixed effects estimation shown in Table 4 is infeasible for quantile regressions since quantiles are not linear operators. Thus, we adopt an estimator that is consistent and asymptotically normal [20] to compute the quantile estimates.

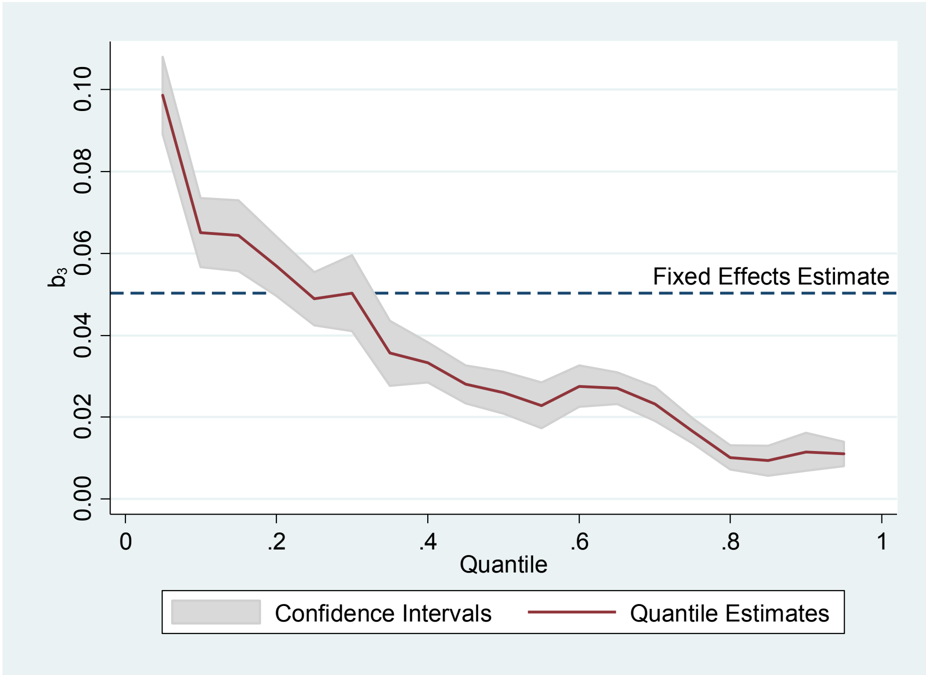

Figure 3. The Estimates of Quantile and Fix Effects Regressions

In Figure 3, we plot the parameter estimates b_3 of the quantile regressions based on equation (1). There are 19 estimated quantile regressions with 0.05, 0.10, …, and 0.95 quantiles. The parameter estimates of the quantile regressions are connected by the solid line, with the shaded area being their 95% confidence intervals. The fixed effects estimate b_3 shown in column 1 of Table 4 is plotted as a horizontal dashed line in this figure. We find that the quantile regression parameter estimates are significantly positive for all quantiles and decrease with quantiles in general. This suggests that the impact of a positive surprise is higher for less popular videos, because the same magnitude of positive surprises implies more pronounced social learning for less popular videos than for popular videos. Since most of popular videos are published by top content creators on YouTube and consumers have high ex-ante expectations on these videos, the impact of an additional positive surprise (social learning) is relatively small. For example, financed by venture capitalists and grants from YouTube, Maker studios produced a popular sketch comedy show called "AsKassem" for YouTube. Consumers who have watched AsKassem #1 - #70 form high expectations on the quality of the new episode #71. Figure 3 also indicates that the panel data model with fixed effects tends to underestimate the impact of a positive surprise for less popular videos and tends to overestimate the impact for popular videos.

A Test of Network Effects

In this section, we test the existence of network effects by testing whether consumers respond to disconfirmation that is irrelevant to video quality (Hypothesis 2). We use the presence of in-stream ads on the first day as a source of exogenous variation (instrument variable) for existing levels of video views. YouTube in-stream ads run only on partner videos during our data period. Only successful content creators are qualified for the partner program, and videos published by them might contain in-stream ads. Note that all of our sample videos are published by top providers who have partnership status. If only learning exists, the negative disconfirmation would have no significant impact on future views because it does not reflect the quality of video content.[5] However, if network effects also exist, the lower-than-expected views caused by an in-stream ad can still reduce future video views and trigger the feedback loop. In other words, the lower-than-expected views mean that a viewer's colleagues and friends are less likely to watch this video. Meanwhile, the viewer has weaker incentives to watch it because this video is less likely to be discussed in social encounters.

For a first day in-stream ad to be a valid instrument variable (IV) for surprise in Model (1), it has to be (i) correlated with the surprise indicator D_j[6] and (ii) uncorrelated with the error term $\mu_{j,t}$, so that an in-stream ad influences video views only through disconfirmation/surprise. We test condition (i) by regressing in-stream ads on the first day ($adDay1_j$) on D_j with heteroskedasticity-robust standard errors, and find the coefficient on D_j is significant. Although we cannot test condition (ii) directly, first day in-stream ads are arguably independent of unobserved factors for video views later on, because all of our sample videos are published by YouTube partners (high quality YouTube channels).

---

[5] In other words, if learning plays a dominate role in the diffusion, the negative disconfirmation will not significantly affect the word of mouth process on video content. For instance, when one's friends recommend a new trending video, they are more likely to talk about the video content instead of the in-stream ads.

[6] We also do a robustness check when D_j is defined as the original disconfirmation residual instead of a dummy variable, and the basic results remain similar.

In our regression model, we have already controlled for the time-invariant video level fixed effect a_j, so there is less of a concern that any potential omitted variables in the error term $\mu_{j,t}$ would be correlated with the first day in-stream ad. Moreover, the inclusion of an in-stream ad is likely to cause negative disconfirmation because consumers may switch to other videos due to annoying ads. In our specific context, we assume that the first day in-stream ad could be treated as excluded from the equation (1). The major concern with this exclusion restriction is that if the presence of in-stream ads influences future growth of viewership other than through disconfirmation, our approach is called into question. Following Angrist and Krueger [3], and Acemoglu et al. [1], we conduct a test by including the term, $(t \times adDay1_j)$, as an independent variable in equation (1). The intuition is that assuming the only impact of in-stream ads on future growth of viewership is through disconfirmation, then the in-stream ad shocks should be insignificant in equation (1) that also includes disconfirmation. We find that when $(t \times adDay1_j)$ is entered as a regressor in equation (1), it has a t-statistic less than 1 after controlling for the interaction term $t \times D_j$. It means that the future growth of viewership is not influenced by the presence of in-stream ads other than through disconfirmation. The reason is that in reality a viewer knows whether there is an in-stream ad only after she starts watching a video.

However, even if condition (i) is satisfied, if the instrument is weak that $adDay1_j$ is only weakly correlated with D_j, IV methods can be ill-behaved so that seemingly very small correlation between the IV and the error term can cause severe inconsistency and therefore severe finite sample bias [63, 2]. For this concern, we test whether $adDay1_j$ is a weak instrument, by calculating the first-stage F statistic based on the method proposed by Stock et al. [54]. For an instrument to be reliable, the first-stage F statistic in the two-stage least squares (2SLS) regression should be greater than 10. We examine F statistics from the first-stage regressions in 2SLS and find all of them are greater than 10.

Our purpose of using instrument variable is to isolate the surprises caused solely by the presence of in-stream ads, and to test whether the surprises specifically driven by ads would influence video views. The presence of in-stream ads is negative disconfirmation that can reduce first-day video views. If network effects exist on YouTube, the negative disconfirmation further lowers the views at time 2. As time goes on, we should see a significantly negative self-reinforcing feedback loop. However, such negative disconfirmation does not reflect any information of the video quality. If learning is the sole form of social contagion, we would not see a self-reinforcing feedback loop.

We re-estimate model (1) to test for network effects, using 2SLS regression with the in-stream ads on the first day, $adDay1_j$, as instrument for the surprise dummy D_j. $adDay1_j$ is a dummy, where $adDay1_j = 1$ if the video has an in-stream ad, and $adDay1_j = 0$ otherwise. Generally, 2SLS is used to deal with endogeneity. We use the first-stage regression of 2SLS to isolate the disconfirmation caused solely by the shock of in-stream ads. We are interested in the IV estimator b_3 on $(t \times D_j)$, in the regression model (1). If the presence of network effects is supported, we expect that $b_3 > 0$, which implies that a video that has positive non-quality disconfirmation would have a higher growth rate of viewership than a video that has negative non-quality disconfirmation over time. In other words, a positive b_3 in the IV estimation implies that compared with positive non-quality disconfirmation, negative non-quality disconfirmation will lower the future growth rate of viewership. The coefficient $b_3 = 0$ would suggest insignificant network effects. The results are shown in column 1 of Table 7. We find that $\hat{b}_3$ is significantly positive, supporting the presence of network effects on YouTube. Column 2 shows that the results still hold when the disconfirmation is defined as the difference between realized video ratings and predicted video ratings. Therefore, Hypothesis 2 is supported. In order to further address the concern of the endogenous instrument variable, we present the estimation results using another instrument variable, the weather on video publish days, in the online appendix.

Table 7. A Test of Network Effects: 2SLS

	(1) IV	(2) Rating as Surprise
$(\boldsymbol{t} \times \boldsymbol{D_j})$	0.0154***	0.00864***
	[25.62]	[12.45]
$\boldsymbol{t}$	0.0312***	0.0216***
	[11.65]	[10.32]
rating	0.292*	0.128
	[1.655]	[0.525]
comment	0.000725***	0.000228***
	[15.40]	[6.872]
favs	5.22e-05	2.48e-05
	[0.799]	[0.524]
sum_upload	0.559***	0.318***
	[4.244]	[3.534]
Observations	9060	9060
R-squared	0.248	0.226

t-statistics in brackets, *** p<0.01, ** p<0.05, * p<0.1

Prior econometrics literature showed that even if instrumental variables do not perfectly satisfy the exclusion restriction, we can still draw valid statistical inferences using the Anderson and Rubin (AR) test and the fractionally resampled Anderson Rubin (FAR) test [49]. As a robustness check, we further conduct these two tests in our IV regression and find that the *p*-values of the coefficient on the interaction term, $t \times D_j$, are less than 0.01, which further confirms our Hypothesis 2.

Application: How to Go Viral?

We have shown that both learning and network effects exist on YouTube. To further test the two effects on high-quality videos and attention-grabbing videos respectively (Hypothesis 3), we estimate the following model:

$$\begin{aligned} \ln views_{jt} &= b_0 + b_1 t + a_j + \Psi'_{jt} b_2 + b_3 (t \times D_j) \\ &+ b_4 (t \times attention_j) + b_5 (t \times quality_j) + b_6 (t \times D_j \times attention_j) \\ &+ b_7 (t \times D_j \times quality_j) + \mu_{jt}, \end{aligned} \tag{2}$$

where binary variable $attention_j$ indicates whether video j is an attention-grabbing video, and binary variable $quality_j$ indicates whether video j is a high-quality video. Here, we use views and rating to empirically identify high-quality videos and attention grabbers. We define high-quality videos as those with many views and high ratings, and attention-grabbing videos as those with many views but mixed ratings. The co-existence of extremely high and extremely low ratings often suggests controversy. Specifically, if both the number of views and the rating of video j at the end of our sample period rank among the top 25% of all sample videos, we consider it to be a high-quality video and $quality_j = 1$; otherwise, $quality_j = 0$. If the number of views of video j at the end of our sample period rank among the top 25%, but the rating is in the lowest 25%, then we consider it to be a video with controversial content, and $attention_j = 1$; otherwise, $attention_j = 0$.

Looking at the growth of views for a typical attention-grabbing video and a typical high-quality video as shown in Figure 4, we find that generally, for the attention-grabbing videos, views increase exponentially in the first week but only marginally thereafter, whereas for high-quality videos, views exhibit a gradual but steady growth over a much longer period. To study learning and network effects for the two types of popular videos, we estimate the regression models (2) and (3) using 2SLS with an instrument variable and fixed effect model without an instrument variable. In 2SLS, we instrument the surprise dummy D_j using the in-stream ads $adDay1_j$.

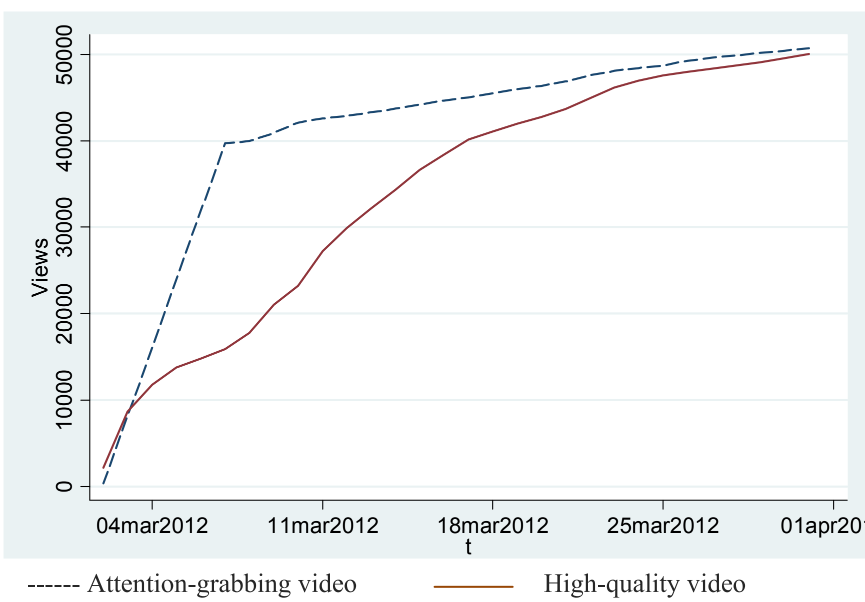

Figure 4. Growth of views for attention-grabbing video and high-quality video

The results are shown in Table 8. Column 1 reports the results from 2SLS estimation of equation (2) with IV. Column 2 reports the results from the fixed effect estimation of equation (2) without IV. We are interested in the coefficients b_6 and b_7 in equation (2). The estimation results show that network effects are significantly stronger for attention-grabbing videos but weaker for high-quality videos; the upper bound of learning is significantly stronger for high-quality videos. Therefore, we can conclude that learning is more pronounced for high-quality videos, whereas network effects are more pronounced for attention-grabbing videos, supporting Hypothesis 3.

Our result suggests that videos will be more likely to go viral through network effects if they provoke controversy and stir heated discussion. This result is consistent with some experimental evidence: Content that evokes high-arousal emotions (e.g., awe, anger, and anxiety) is more viral [15]. This finding can help YouTube providers craft contagious content and produce viral videos. Our study also provides empirical support for the strategic use of attention grabbers [25].

Table 8. High-Quality Videos vs. Attention Grabbers

	(1) IV	(2) FE Without IV
$\boldsymbol{t \times D_j \times attention_j}$	0.0105***	0.0119***
	[3.919]	[5.119]
$\boldsymbol{t \times D_j \times quality_j}$	-0.125	0.0375***
	[-1.059]	[4.226]
$(\boldsymbol{t \times D_j})$	0.0702***	0.0494***
	[19.91]	[31.93]
$\boldsymbol{t \times attention_j}$	0.0174	0.0276
	[0.0688]	[1.394]
$\boldsymbol{t \times quality_j}$	0.0410	0.00589
	[1.106]	[1.606]
$\boldsymbol{t}$	0.427**	0.416**
	[2.016]	[2.024]
rating	0.147	0.184
	[0.762]	[0.108]
comment	0.000721***	0.000854***
	[9.564]	[19.02]
favs	2.11e-05	-4.59e-05
	[0.261]	[-0.708]
sum_upload	0.596***	0.552***
	[4.362]	[4.291]

Observations	9060	9060
R-squared	0.223	0.260

t-statistics in brackets, *** p<0.01, ** p<0.05, * p<0.1

Conclusions

In this study, we examine the role of learning and network effects in the diffusion of social media content. Using data from YouTube, we identified learning by (1) estimating the overall effect of first day disconfirmation on the subsequent views, and (2) examining how this effect varies with different prior precisions. Network effects were measured by estimating the effect of the first day disconfirmation that is caused by in-video advertisement and unrelated to video quality, on the subsequent views. We quantified the magnitude of learning and network effects and found that social media content consumption is affected by both learning and network effects.

A straightforward implication of our study is that YouTube should take learning and network effects into account when promoting the growth of video views. As the influence of YouTube on our society, education, entertainment, and lifestyle increases, more and more organizations, including government agencies, TV networks, commercial companies, universities, and so on, are all seeking their own presence and influence in social media. How to manage the influence of consumer buzz in social media is a challenge for practitioners [29, 37]. Our findings provide valuable insights on how to achieve this objective with videos on YouTube.

From a managerial perspective, YouTube can play a much greater role in encouraging the creation of original content by leveraging the multiplier effect of both learning and network effects. In fact, YouTube has nurtured and subsidized individual content creation since its beginning. Tang et al. [58] showed that video providers are indeed encouraged by both reputation and monetary income to produce videos on YouTube. However, whether the average quality of the videos goes up as the quantity increases is still doubtful. Our findings suggested that instead of quality improvement, many providers are generating viral videos leveraging network effects. Although these videos also attract views, how sustainable they are for video providers and YouTube in the long run is still questionable.

Although in the present study we focused only on learning and network effects on UGC sites, our tests are relatively generalizable and can be practically carried out by practitioners in social media. We categorized the most popular videos on YouTube into quality-oriented videos and attention-grabbing videos, and found that videos with attention-grabbing content initiate higher network effects than quality-oriented productions. These findings provide a nuanced view of how YouTube providers can produce viral videos.

Although this study has highlighted the importance of learning and network effects, we do not have social network data, and our work does not consider the effect of network characteristics and network topological structure on social contagion [28]. Further work could incorporate network data and Google Trends data on the keywords of the video to examine the effect of network structure, tie strength, and public opinions on consumer learning and network effects. It would also be interesting to examine under what conditions low-quality YouTube videos may go viral because of learning or network effects.

Acknowledgments: The authors would like to thank Jesse Shore, Gal Oestreicher-Singer, Shun-Yang Lee and the participants of the 2013 Workshop on Information Systems and Economics (WISE) for the helpful feedback. The authors are grateful to the editor and the three anonymous reviewers for their constructive comments that have led to the significant improvements of this paper.

References

1. Acemoglu, D., Johnson, S., and Robinson, J. A. 2001. “The Colonial Origins of Comparative Development: An Empirical Investigation,” *American Economic Review*, 91(5), 1369-1401.
2. Angrist, J. D., and Pischke, J. S. 2008. “Mostly harmless econometrics: An empiricist's companion,”

Princeton university press.

3. Angrist, J. D., and Krueger, A. B. 1991. “Does Compulsory School Attendance Affect Schooling and Earnings?.” *Quarterly Journal of Economics*, 106(4), 979-1014.
4. Angrist, J. D., and Krueger, A. B. 2001. “Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments,” *Journal of Economic Perspectives*, 15(4), 69-85.
5. Amemiya, T. 1973. “Regression Analysis When the Dependent Variable is Truncated Normal,” *Econometrica*, 41(6), 997–1016.
6. Anderson, E. W. 1998. “Customer Satisfaction and Word of Mouth,” *Journal of Service Research*, 1(1), 5-17.
7. Aral, S., Muchnik, L. and Sundararajan, A. 2009. “Distinguish Influence-Based Contagion from Homophily-Driven Diffusion in Dynamic Networks,” *Proceedings of the National Academy of Sciences* (106)51, pp. 21544-21549.
8. Aral, S., and Walker D. 2011. “Creating Social Contagion through Viral Product Design: A Randomized Trial of Peer Influence in Networks,” *Management Science* (57)9, pp. 1623-1639.
9. Arellano, M., and Bond, S. 1991. “Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations,” *Review of Economic Studies*, 58(2), 277-297.
10. Arellano, M., and Bover, O. 1995. “Another look at the instrumental variable estimation of error-components models,” *Journal of Econometrics*, 68(1), 29-51.
11. Baillie, R. T. 1987. “Inference in Dynamic Models Containing Surprise Variables,” *Journal of Econometrics*, (35)1, pp. 101-117.
12. Banerjee, A. V. 1992. “A Simple Model of Herd Behavior,” *Quarterly Journal of Economics*, 107(3), 797-817.
13. Bapna, R., and Umyarov, A. 2014. “Do Your Online Friends Make You Pay? A Randomized Field

Experiment in an Online Music Social Network," *Management Science*, forthcoming.

14. Barro, R. J. 1977. "Unanticipated Money Growth and Unemployment in the United States," *American Economic Review*, (67)2, pp. 101-115.
15. Berger, J., and Milkman, K. L. 2012. "What Makes Online Content Viral?" *Journal of Marketing Research* (49)2, pp. 192-205.
16. Bikhchandani, S., Hirshleifer, D., and Welch, I. 1998. "Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades," *Journal of Economic Perspectives*, 12(3), 151-170.
17. Blundell, R., and Bond, S. 1998. "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, 87(1), 115-143.
18. Bresnahan, T. 2004. "Network Effects and Microsoft," Stanford University Working Paper.
19. Cai, H., Chen Y., and Fang H. 2009. "Observational Learning: Evidence from a Randomized Natural Field Experiment," *American Economic Review* (99)3, pp. 864-882.
20. Cameron, A. C., and Trivedi, P. K.2005. *Microeconometrics: Methods and Applications*, Cambridge University Press.
21. Choi, J. 1997. "Herd Behavior, the 'Penguin Effect', and the Suppression of Informational Diffusion: An Analysis of Informational Externalities and Payoff Interdependency," *Rand Journal of Economics*, 28(3), pp. 407-425.
22. Church, J., and Gandal, N. 1993. "Complementary Network Externalities and Technological Adoption," *International Journal of Industrial Organization*, 11(2), pp. 239-260.
23. Drehmann, M., Oechssler, J., and Roider, A., 2007. "Herding with and without Payoff Externalities: An Internet Experiment," *International Journal of Industrial Organization*, 25(2), pp. 391–415.
24. Economides, N. 1996. "The Economics of Networks," *International Journal of Industrial Organization*, 14(6), pp. 673-699.
25. Eliaz, K., and Spiegler, R. 2011, "On the Strategic Use of Attention Grabbers," *Theoretical*

Economics (6)1, pp. 127-155.

26. Gale, D. 1996. "What Have We Learned from Social Learning?" *European Economic Review*, 40(3-5), pp. 617-628.
27. Gallaugher, J. M., and Wang, Y. 2002. "Understanding Network Effects in Software Markets: Evidence from Web Service Pricing," *Management Information Systems Quarterly*, 26(4), pp. 303-327.
28. Ghose, A., Han, S., and Iyengar, R.2012. "An Empirical Analysis of Network Characteristics and Social Contagion on the Mobile Internet," working paper.
29. Gu, B., and Ye, Q. 2014. "First Step in Social Media: Measuring the Influence of Online Management Responses on Customer Satisfaction," *Production and Operations Management*, 23(4), 570-582.
30. Hirshleifer, D., Lim, S. S., and Teoh, S. H. 2009. "Driven to Distraction: Extraneous Events and Underreaction to Earnings News," *Journal of Finance*, (64)5, pp. 2289-2325.
31. Ho, T., Li, S., Park, S. E., and Shen, M. Z. 2012. "Customer Influence Value and Purchase Acceleration in New Product Diffusion," *Marketing Science*, (31)2, pp. 236-256.
32. Hu, Y., and Schennach, S. M. (2008). Instrumental variable treatment of nonclassical measurement error models. *Econometrica*, 76(1), 195-216.
33. Katona, Z., Zubcsek, P. P., and Sarvary, M. 2011. "Network Effects and Personal Influences: The Diffusion of an Online Social Network," *Journal of Marketing Research* (48)3, pp. 425-443.
34. Katz, M. L., and Shapiro, C. 1985. "Network Externalities, Competition and Compatibility," *American Economic Review* (75)3, pp. 424–440.
35. Kitroeff, N. 2014. "Why That Video Went Viral," NYtimes.com. Available at: http://www.nytimes.com/2014/05/20/science/why-that-video-went-viral.html
36. Lee, G., Raghu, T. S. 2014, "Determinants of Mobile Apps Success: Evidence from App Store Market," *Journal of Management Information Systems*, 31(2), pp. 133-170.

37. Luo, X., and Zhang, J. 2013. "How Do Consumer Buzz and Traffic in Social Media Marketing Predict the Value of the Firm?." *Journal of Management Information Systems*, 30(2), 213-238.

38. Manski, C. F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies* (60)3, pp. 531-542.

39. McKinney, V., Yoon, K., and Zahedi, F. M. 2002. "The measurement of web-customer satisfaction: an expectation and disconfirmation approach," Information Systems Research, 13(3), 296-315.

40. Miguel, E., Satyanath, S., and Sergenti, E. (2004). "Economic shocks and civil conflict: An instrumental variables approach," *Journal of Political Economy*,112(4), 725-753.

41. Moretti, E. 2011, "Social Learning and Peer Effects in Consumption: Evidence from Movie Sales," *Review of Economic Studies*, (78)1, pp. 356-393.

42. Muth, J. F. 1961. "Rational Expectations and the Theory of Price Movements." *Econometrica*, 29(3), 315-335.

43. Olson, J. C., and Dover, P. A. 1979. Disconfirmation of consumer expectations through product trial. *Journal of Applied Psychology*, 64(2), pp. 179-189.

44. Oliver, R. L. 1980. A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research*, (17)4, pp. 460-469.

45. Pagan, A. 1984. "Econometric Issues in the Analysis of Regressions with Generated Regressors," *International Economic Review*, (25)1, pp. 221-247.

46. Qiu, L., and Whinston A. B. 2012. "Naive Herding in Location-Based Networks: Induced Social Learning and Optimal Dynamic Pricing," working paper, Available at SSRN: http://ssrn.com/abstract=2090093.

47. Qiu, L., Rui, H., and Whinston, A. B. 2014. "Effects of Social Networks on Prediction Markets: Examination in a Controlled Experiment," *Journal of Management Information Systems*, 30(4), 235-268.

48. Richins, M. L. 1983. "Negative word-of-mouth by dissatisfied consumers: a pilot study," *Journal*

of Marketing, 47(1), pp. 68-78.

49. Riquelme, A., Berkowitz, D., and Caner, M. (2013). "Valid tests when instrumental variables do not perfectly satisfy the exclusion restriction," *Stata Journal*, 13(3), 528-546.
50. Sargent, T. J. 1976. "A Classical Macroeconometric Model for the United States," *Journal of Political Economy* (84)2, pp. 207-237.
51. Scharfstein, D., and Stein, J. 1990. "Herd Behavior and Investment," A*merican Economic Review*, 80(3), 465-479.
52. Shi, Z., and Whinston, A. B. 2013. "Network Structure and Observational Learning: Evidence from a Location-Based Social Network," *Journal of Management Information Systems*, 30(2), 185-212.
53. Smith, L., and Sorensen, P. 2000. "Pathological Outcomes of Observational Learning," *Econometrica*, 68(2), pp. 371-398.
54. Stock, J. H., Wright, J. H., and Yogo, M. 2002. "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business & Economic Statistics* 20(4), pp. 518-529
55. Susarla, A., Barua, A., and Whinston, A. B. 2003. Understanding the service component of application service provision: empirical analysis of satisfaction with ASP services. *Management Information Systems Quarterly*, 27(1), 91-123.
56. Susarla, A., Oh, J., and Tan, Y, 2012. "Social Networks and the Diffusion of User-Generated Content: Evidence from YouTube," *Information Systems Research* 23(1), pp. 23-41.
57. Susarla, A., Tan, Y., and Oh, J. 2013. Influentials, Imitables or Susceptibles? Virality and Word-of-Mouth Conversations in Online Social Networks. Available at SSRN: http://ssrn.com/abstract=1984690.
58. Tang, Q., Gu, B., and Whinston, A. B. 2012. "Content Contribution for Revenue Sharing and Reputation in Social Media: A Dynamic Structural Model," *Journal of Management Information Systems* (29)2, pp. 41-75.

59. van Osselaer, S. J. and Alba, J. W. 2000. “Consumer Learning and Brand Equity,” *Journal of Consumer Research*, 27 (1), 1–16.

60. Walker, R. 2010. “When Funny Goes Viral,” NYtimes.com. Available at: http://www.nytimes.com/2010/07/18/magazine/18ROFL-t.html

61. Wang, A., Zhang, M., and Hann, I. H. 2014. “Socially Nudged: A Quasi-Experimental Study of Friends' Social Influence in Online Product Ratings,” *Information Systems Research*, forthcoming.

62. Williams, S., and Radner, R. 1988. "Informational Externalities and the Scope of Efficient Dominant Strategy Mechanisms," *Discussion Paper #761*, Northwestern University.

63. Wooldridge, J.M. 2002. *Econometrics Analysis of Cross Section and Panel Data*, MIT Press: Cambridge, MA.

64. Zhang, X., and Wang, C. 2012. “Network positions and contributions to online public goods: The case of Chinese Wikipedia,” *Journal of Management Information Systems*, 29(2), 11-40.

Appendix A

Result 1: If positive/negative disconfirmation is sufficiently large, then Π_{iT} is increasing/decreasing in T.

Proof:

$$\Pi_{iT+1} - \Pi_{iT}$$

$$= \Phi\left(\frac{\alpha_{T+1}X_j'\beta + (1-\alpha_{T+1})V_j - c_{ij}}{g_{T+1}}\right) - \Phi\left(\frac{\alpha_T X_j'\beta + (1-\alpha_T)V_j - c_{ij}}{g_T}\right).$$

We also have

$$\frac{\alpha_{T+1}X_j'\beta + (1-\alpha_{T+1})V_j - c_{ij}}{g_{T+1}} - \frac{\alpha_T X_j'\beta + (1-\alpha_T)V_j - c_{ij}}{g_T}$$

$$= \xi_T\left(V_j - X_j'\beta\right) - \left(\frac{1}{g_T} - \frac{1}{g_{T+1}}\right)\left(X_j'\beta - c_{ij}\right),$$

where $\xi_T = \frac{1-\alpha_{T+1}}{g_{T+1}} - \frac{1-\alpha_T}{g_T}$

$$= \frac{1}{\sqrt{1/(\sum_{t=2}^{T+1} d_{it}) + (1/\rho_\eta)\left(\rho_{V_j}/\sum_{t=2}^{T+1} d_{it} + 1\right)^2}} - \frac{1}{\sqrt{1/(\sum_{t=2}^{T} d_{it}) + (1/\rho_\eta)\left(\rho_{V_j}/\sum_{t=2}^{T} d_{it} + 1\right)^2}} > 0.$$

Since $V_j \gg X_j'\beta$ and $\xi_T > 0$, we can let V_j be sufficiently large, such as $V_j > \frac{1}{\xi_T}\left(\frac{1}{g_T} - \frac{1}{g_{T+1}}\right)\left(X_j'\beta - c_{ij}\right) + X_j'\beta$. Therefore, We can obtain $\Pi_{iT+1} - \Pi_{iT} > 0$. Thus, if a positive surprise is sufficiently large, then Π_{iT} is increasing in T. Similarly, we can show that if a negative surprise is sufficiently large, then Π_{iT} is decreasing in T. ■

Result 2: The distance between the true video quality and consumers' expected video quality, $|V_j - E_T[V_j|I_T]|$, is decreasing in T.

Proof: $|V_j - E_T[V_j|I_T]| = |V_j - \alpha_T X_j'\beta - (1-\alpha_T)V_j| = \alpha_T|V_j - X_j'\beta|$. $\alpha_T = \frac{\rho_{V_j}}{\rho_{V_j} + \sum_{t=2}^{T} d_{ijt}}$ and $d_{ijt} > 0$, so α_T is decreasing in T. Therefore, $\alpha_T|V_j - X_j'\beta|$ decreases with T. ■

Appendix B

Table B.1. Correlations among Variables in First-Stage Regression

	lvviews	lcviews	lvideos	lsubs	lsubscriptions	lchannel_favs	length	avg_chanel_rating	var_chanel_rating	channel_age	tweet_upload
lvviews	1										
lcviews	0.6699	1									
lvideos	-0.0188	0.0411	1								
lsubs	0.4726	0.4550	-0.1135	1							
lsubscriptions	0.0715	0.1566	-0.0278	0.1157	1						
lchannel_favs	0.0716	0.1501	-0.0563	0.1851	0.4643	1					
length	-0.0592	0.0549	-0.1596	-0.0132	0.0218	-0.0385	1				
avg_chanel_rating	0.5232	0.3591	0.0022	0.3282	0.1257	0.1528	0.0079	1			
var_chanel_rating	-0.1141	-0.0899	0.2136	-0.1846	-0.1129	-0.1575	-0.0065	-0.2115	1		
channel_age	0.0197	-0.0191	-0.0405	0.0420	-0.1125	0.0197	-0.2830	-0.0294	0.0402	1	
tweet_upload	-0.0180	-0.0211	-0.0197	0.0271	0.0622	0.0102	-0.0208	-0.0221	-0.0459	-0.0744	1

Table B.2. Variance Inflation Factor of First-Stage Regression

	VIF
lvviews	2.63
lcviews	2.37
lvideos	1.28
lsubs	1.49
lsubscriptions	1.51
lchannel_favs	1.60
length	1.47
avg_chanel_rating	1.51
var_chanel_rating	1.20
channel_age	1.43
tweet_upload	1.06

Note: The small VIF values show that the multicollinearity problem is not serious